·论著·

胃癌发生风险的列线图预测模型研究

周倩[1]，吴晓敏[1，2]，王宝华[1]，严若菡[1，2]，蔚苗[1，3]，吴静[1*]

1.100050 北京，中国疾病预防控制中心慢性非传染性疾病预防控制中心
2.300070 天津，天津医科大学公共卫生学院
3.010107 内蒙古呼和浩特，内蒙古医科大学公共卫生学院
*通信作者：吴静，研究员 / 博士生导师；E-mail：wujing@chinacdc.cn

【摘要】 **背景** 胃癌（Gastric cancer，GC）严重危害我国居民健康，对个体 GC 发病风险进行预测有助于早期识别高风险人群，进而采取有针对性的干预措施避免或延缓 GC 进展。**目的** 构建并验证预测个体 GC 发生风险的列线图模型。**方法** 选取 2020 年 1 月—2021 年 7 月在安徽、河南、山东、江苏四省 14 个县（区）肿瘤登记系统中确诊的≥ 40 岁的 GC 患者作为病例组（684 例），按 1∶2 比例纳入性别、年龄、居住地及健康状况匹配的一般人群作为对照组（1368 例）。将所有研究对象按 8∶2 比例随机分为训练集（1 641 例）和验证集（411 例）。采用多因素 Logistic 回归分析筛选变量，建立列线图预测模型。绘制模型预测 GC 发病风险的受试者工作特征（ROC）曲线，采用 ROC 曲线下面积（AUC）、Hosmer-Lemeshow 法评估模型的区分度和校准度；采用 Bootstrap 法进行模型验证，采用临床决策曲线（DCA）评估模型的临床实用性。**结果** 多因素 Logistic 回归分析结果显示，饮食口味偏咸（*OR*=1.690，95%*CI*=1.333~2.142）、食物偏硬（*OR*=1.596，95%*CI*=1.145~2.225）、喜食辣食（*OR*=1.387，95%*CI*=1.093~1.760）、二手烟暴露（*OR*=1.880，95%*CI*=1.473~2.399）、经常发脾气（*OR*=3.283，95%*CI*=2.236~4.819）、胃部疾病史（*OR*=4.008，95%*CI*=3.046~5.273）、一级亲属肿瘤史（*OR*=1.549，95%*CI*=1.170~2.051）、幽门螺杆菌感染（*OR*=1.298，95%*CI*=1.028~1.693）、高盐饮食（*OR*=1.338，95%*CI*=1.033~1.734）是 GC 发生的独立危险因素（$P<0.05$）；高中及以上学历（*OR*=0.491，95%*CI*=0.342~0.703）、规律饮食（*OR*=0.542，95%*CI*=0.405~0.726）、食生蒜或蒜苗（*OR*=0.501，95%*CI*=0.394~0.636）是 GC 发生的保护因素（$P<0.05$）。训练集和验证集预测 GC 发生风险的 ROC 曲线下面积（AUC）分别为 0.768（95%*CI*=0.744~0.792）和 0.776（95%*CI*=0.728~0.823）。Bootstrap 法验证结果显示，校正曲线与实际曲线一致性良好（训练集：Brier 评分 =0.177；验证集：Brier 评分 =0.176）；Hosmer-Lemeshow 检验结果显示，模型拟合度良好（训练集：χ^2=4.408，P=0.819；验证集：χ^2=4.650，P=0.794）。DCA 曲线显示当阈值为 0.05~0.79 时，使用列线图模型预测 GC 发生风险可以使患者临床获益。**结论** 本研究构建的列线图模型可预测 GC 的发病风险，便于早期识别高风险人群，且有助于制定有针对性的个体化干预措施。

【关键词】 胃癌；危险因素；列线图；预测模型；病例对照研究；Logistic 回归
【中图分类号】 R 735.2 【文献标识码】 A DOI：10.12114/j.issn.1007-9572.2024.0525

Study on Nomogram Prediction Model for Risk of Gastric Cancer

ZHOU Qian1，WU Xiaomin1，2，WANG Baohua1，YAN Ruohan1，2，YU Miao1，3，WU Jing1*
1.National Center for Chronic and Non-Communicable Disease Control and Prevention，Chinese Center for Disease Control and Prevention，Beijing 100050，China
2.School of Public Health，Tianjin Medical University，Tianjin 300070，China
3.School of Public Health，Inner Mongolia Medical University，Hohhot 010107，China
**Corresponding author：WU Jing，Researcher/Doctoral supervisor；E-mail：wujing@chinacdc.cn*

【Abstract】 **Background** Gastric cancer（GC）causes a heavy burden in China. Predicting individual GC risk can help to identify high-risk groups early，and then take targeted interventions to avoid or delay GC progression. **Objective** Establish and validate a nomogram model for predicting individual GC risk. **Methods** From January 2020 to July 2021，

引用本文：周倩，吴晓敏，王宝华，等. 胃癌发生风险的列线图预测模型研究［J］. 中国全科医学，2024. DOI：10.12114/j.issn.1007-9572.2024.0525.［Epub ahead of print］［www.chinagp.net］
ZHOU Q，WU X M，WANG B H，et al. Study on nomogram prediction model for risk of gastric cancer［J］. Chinese General Practice，2024.［Epub ahead of print］



GC patients ≥ 40 years were diagnosed from the cancer registry system in 14 counties (districts) of Anhui province, Henan province, Shandong province and Jiangsu province were selected as the case group (684 cases). Match the general population with a frequency of 1 : 2 based on gender, age, place of residence and health status matching people as a control group (1 368 cases). All subjects were randomly divided into training set (1 641 cases) and validation set (411 cases) according to a ratio of 8: 2. Multivariate Logistic regression analysis was used to screen variables and establish nomogram prediction model. The receiver operating characteristic (ROC) curve of the model predicting the risk of GC was drawn, the discrimination and calibration of the model were evaluated via the area under the ROC curve (AUC) and the Hosmer-Lemeshow test. The model was verified by Bootstrap method and the decision curve analysis (DCA) was used to evaluate the clinical practicability of the model. **Results** Multivariate Logistic regression analysis showed that salty tastes (*OR*=1.690, 95%*CI*=1.333-2.142), dry and hard diet (*OR*=1.596, 95%*CI*=1.145-2.225), spicy food tastes (*OR*=1.387, 95%*CI*=1.093-1.760), exposure to secondhand smoking (*OR*=1.880, 95%*CI*=1.473-2.399), frequent tantrums (*OR*=3.283, 95%*CI*=2.236-4.819), history of stomach disease (*OR*=4.008, 95%*CI*=3.046-5.273), the family history of cancer (*OR*=1.549, 95%*CI*=1.170-2.051), Helicobacter pylori (Hp) infection (*OR*=1.298, 95%*CI*=1.028-1.693), high-salt diet(*OR*=1.338, 95%*CI*=1.033-1.734) were independent risk factors for GC(*P*<0.05). Junior high school education(*OR*=0.616, 95%*CI*=0.4688-0.811), high school education or above (*OR*=0.491, 95%*CI*=0.342-0.703), regular diet (*OR*=0.542, 95%*CI*=0.405-0.726), the garlic consumption (*OR*=0.501, 95%*CI*=0.394-0.636) were protective factors for GC (*P*<0.05). The area under ROC curve (AUC) for predicting GC risk in the training and validation sets was 0.768 (95%*CI*=0.744-0.792) and 0.776 (95%*CI*=0.728-0.823), respectively. The verification results of Bootstrap method showed that the calibration curve was in good agreement with the actual curve (Brier score of training set=0.177; Brier score of verification set=0.176); Hosmer-Lemeshow results showed that the model had a good fit (Training set: χ^2=4.408, *P*=0.8186; Verification set: χ^2=4.650, *P*=0.7943). Bootstrap results showed that the model is in good agreement. The DCA curve showed that when the threshold is between 0.05 and 0.79, patients can benefit clinically using the nomogram model to predict the risk of GC occurrence. **Conclusion** The nomogram model constructed in this study could predict individual GC risk, early identify high-risk groups and help to formulate targeted and individualized interventions.

【Key words】 Gastric cancer; Risk factors; Nomogram; Prediction model; Case-control study; Logistic regression

胃癌（Gastric cancer，GC）严重威胁人类健康。数据显示，2020 年全球新增 108.9 万 GC 病例，76.9 万死亡病例（全球每 13 例中死亡 1 例）[1]。2020 年我国 GC 新发病例数为 47.9 万，死亡病例数达 37.4 万，分别占全球的 43.94% 和 48.62%，居恶性肿瘤第 3 位[2]。我国 GC 发病率和死亡率在全球范围内仍居高位，疾病负担沉重。

研究表明，早期 GC 手术治疗后 5 年生存率可达 97.6%[3]，进展期 GC 积极治疗后，其 5 年生存率仅为 35% 左右[4]。且早期 GC 症状不明显，多数患者首诊时已处于进展期，早期发现和治疗对降低 GC 发病和死亡至关重要。内镜及活检是目前诊断 GC 的金标准，但其属于侵入性诊断方法，费用相对较高且患者接受程度差[5]，不适用于整个人群的 GC 筛查。应用非侵入性诊断方法作为初始预筛查工具，预筛选后再进行内镜精查是更可行的筛查策略。

近年来，风险预测模型广泛应用于癌症防治领域中，肿瘤风险预测模型能定量识别高危人群的患癌风险，有效降低癌症的发病和死亡[6-7]。因此，本研究基于饮食习惯、生活方式、胃部疾病史、幽门螺杆菌（Helicobacter pylori，*Hp*）感染等 GC 危险因素，构建定量、直观的 GC 风险预测模型，并对模型效能、稳定性和临床效用进行评估，以作为初始筛查工具准确识别高风险人群，从而为 GC 早期预防和治疗提供参考依据。

1 对象与方法

1.1 研究对象

本研究为 1：2 匹配病例对照研究。选取安徽、河南、山东、江苏 4 省中 14 个县（区）（包括：江苏省盱眙县、金湖县和射阳县；安徽省颍东区、埇桥区、灵璧县、蒙城县和寿县；山东省汶上县、巨野县；河南省扶沟县、西平县、罗山县和沈丘县，具体入组人数见附表 1）肿瘤登记系统中于 2020 年 1 月—2021 年 7 月确诊的≥ 40 岁 GC 患者作为病例组，选取性别、年龄、居住地及健康状况符合纳入标准的一般人群作为对照组。

1.1.1 样本量计算：通过本次调查的研究因素在人群中的暴露率（P_0）和与疾病关联强度的估计值（OR）以及统计学检验假设的显著性水平（α）和检验效能（$1-\beta$）估算病例组所需样本量 n，对照组数为 $2n$。选择 Hp 感染因素为主要暴露因素计算样本量。设置检验水准 α=0.025，检验效能 $1-\beta$=0.95，能获得更高的检验水准。根据既往文献可知 *Hp* 感染与胃癌 *OR*=2.56[8]，中国成

人 *Hp* 暴露率 P_0=0.43[9]，危险因素相关程度 Phi=0.2，根据 PASS 软件计算结果得 *n*=123。考虑研究对象的无应答率为 10%，则需样本量为病例组 137 例，对照组 274 例。本研究在计算所得样本量基础上将样本扩大 4 倍，病例组为 548 例，对照组为 1 096 例。通过肿瘤登记系统，在 14 个县（区）各选取胃癌新发病例 50 例，并按要求选取一般人群作为对照。最终共收集 2 139 例调查对象信息，整理清洗数据后，删除 87 例存在数据缺失或无法进行匹配的调查个例，最后共纳入 2 052 例，病例组 684 例，对照组 1 368 例，见附表 1。

1.1.2　纳入排除标准：病例组：（1）纳入标准：GC 患者年龄≥ 40 岁，选取在各县（区）确诊时间在 2020 年 1 月—2021 年 7 月的新发病例；精神状况正常，无认知障碍；在本地居住≥ 10 年；无其他既往恶性肿瘤病史；诊断标准同《2019 年中国肿瘤登记年报》一致，疾病分类编码为 C16；调查开展时存活的胃癌患者。（2）排除标准：年龄 <40 岁；因健康问题等无法参与调查、疾病严重，无法配合或丧失语言或行动能力的受试者；未在本地居住≥ 10 年，或因外出务工等其他原因无法返回当地参与调查的受试者；患其他肿瘤的病人；诊断标准与《2019 年中国肿瘤登记年报》不同者。

对照组：（1）纳入标准：年龄≥ 40 岁且与病例组患者相差 5 岁以内，与病例性别相同者；精神状况正常，无认知障碍；在本地居住≥ 10年；无既往恶性肿瘤病史。（2）排除标准：对照与病例超过 5 岁的年龄差，与病例性别不同者；因健康问题等不能或不配合调查的人；因外出务工，或其他原因无法返回当地参与调查的受试者；有既往恶性肿瘤病史。

本研究未对研究对象实施干预，且通过了中国疾病预防控制中心慢性病与非传染性疾病预防控制中心伦理审查委员会的审查（批准号：202025）。

1.2　研究方法

1.2.1　一般资料采集

采用问卷调查方式收集研究对象基本信息，问卷调查主要包括基本信息（姓名、性别、出生日期、民族、地区、文化程度、居住方式、住址）、饮食习惯、生活方式、性格情绪特征及精神状态、胃部疾病史及家族肿瘤史等。体格检查包括身高和体重测量。饮食习惯包括：饮食口味偏咸（指在日常饮食中倾向于选择含盐量较高或经过咸味加工的食物，并且更喜欢咸味带来的口感和风味，如腌制食品、方便面等方便食品、午餐肉等罐头食品等）、食用食物偏硬（食物选择上偏好质地较硬、不易咀嚼的食物，如肉类干制品、坚果、干果、粗粮等）、喜食辛辣食物（指在食物选择上偏好辛辣口味，即喜欢食用添加了辣椒、芥末、花椒、姜等辛辣调料的食物）、规律饮食（指按照一定的时间间隔及合理的膳食结构来安排每日的饮食，即每天定时、定量和均衡饮食）。胃部疾病史：指在过去患有或现患有胃溃疡、萎缩性胃炎或胃食管反流病一种或几种的记录或历程。

1.2.2　盐阈值检测

采用不同浓度的盐阈值口嚼片进行检测。共有 4 种盐浓度 0.3%、0.6%、0.9% 和 1.2%，分别对应日均盐摄入量 4、6、8、12 g。测试方法：受试者清水漱口后依次咀嚼 4 种盐浓度口嚼片尝试盐味浓度，每次咀嚼 10 次左右，当尝至与自身日常饮食口味最相当的咸度时对应的口嚼片号数记录为盐阈值。盐阈值 >2 则认为研究对象处于高盐饮食状态。

1.2.3　Hp 检测

采用尿素 ^{13}C 呼气诊断试验试剂盒（北京华亘安邦科技有限公司，货号：6655866449047186），进行呼气试验测定。操作方法：空腹 2 h 以上的受试者向 0 min 集气袋吹气，并加盖密封；在试剂瓶中加入饮用水至刻度线，加盖摇晃至试剂充分溶解后让受试者服用；静坐 30 min 后开启 30 min 集气袋，向其中平静吹气至气袋鼓起，加盖密封；将集气袋贴好标签后常温储存，待送检。

问卷调查、体格检查及盐阈值检测均由培训合格的疾病预防控制中心工作人员完成，Hp 检测由试剂检测公司（华亘实验室）的专业检测人员完成。

1.3　统计学方法

采用 SAS 9.4 和 R 4.2.3 软件对数据进行统计分析。符合正态分布的计量资料采用（$\bar{x} \pm s$）表示，组间比较采用独立样本 *t* 检验；不符合正态分布数据以 M（P_{25}，P_{75}）表示，组间比较采用秩和检验；计数资料采用构成比表示，组间比较采用 χ^2 检验。多因素 Logistic 回归筛选独立危险因素，构建列线图模型。采用 Bootstrap 法进行内部验证，绘制受试者工作特征（ROC）曲线，并计算 ROC 曲线下面积（AUC），评估模型区分度。采用 Hosmer-Lemeshow 检验评价模型拟合度。采用临床决策曲线（DCA）评估模型的临床实用性。以 $P<0.05$ 差异具有统计学意义。

2　结果

2.1　一般资料

共纳入 2 052 例研究对象，其中男性 1 479 例（72.08%），女性 573 例（27.92%），年龄 67.0（58.0，72.0）岁。病例组与对照组性别、年龄、民族、居住方式、食奶制品比较，差异无统计学意义（$P>0.05$）。病例组与对照组文化程度、饮食口味偏好、食物软硬偏好、辛辣食物偏好、吃饭时长、规律饮食、畜肉食用量、食豆制品、食生蒜或蒜苗、零食、饮茶、二手烟暴露、锻炼、发脾气、生闷气、胃部疾病史、一级亲属肿瘤史、Hp 感染、高盐饮食比较，差异有统计学意义（$P>0.05$），

见表 1。

2.2　GC 发生风险的多因素 Logistic 回归分析

以是否患 GC（赋值：否 =0，是 =1）为因变量，将表 1 中差异有统计学意义的 19 个变量（文化程度、饮食口味偏好、食物软硬偏好、辛辣食物偏好、吃饭时长、规律饮食、食畜肉、食豆制品、食生蒜或蒜苗、零食、饮茶、二手烟暴露、锻炼、发脾气、生闷气、胃部疾病史、一级亲属肿瘤史、*Hp* 感染、高盐饮食）纳入多因素 Logistic 回归分析，采用逐步回归法筛选变量，变量赋值见表 2。结果显示，饮食口味偏咸、食物偏硬、喜食辛辣食物、二手烟暴露、经常发脾气、胃部疾病史、一级亲属肿瘤史、*Hp* 感染、高盐饮食（盐阈值 >2）是 GC 的危险因素，初中及以上学历、规律饮食、食用生蒜或蒜苗是 GC 的保护因素（P<0.05），见表 3。

2.3　GC 发生风险的列线图模型

纳入多因素 Logistic 回归分析有统计学意义的变量绘制列线图，将各预测变量的分数相加得到总分，通过向下绘制垂线得出 GC 发病的估计概率。以高中及以上学历、饮食口味清淡、食用食物偏软、喜食辣食、规律饮食、食用生蒜或蒜苗、有二手烟暴露、不常发脾气、

表 1　病例组与对照组一般资料比较

Table 1　Comparison of general information between case group and control group

组别	例数	性别［例（%）］		年龄［M（P_{25}，P_{75}）］	民族［例（%）］		居住方式［例（%）］		文化程度［例（%）］		
		男	女		汉族	回族	独居	与配偶同住	小学及以下	初中	高中及以上
例数	1 479	573	2 052	2 041	11	263	1 789	1 192	554	306	776
病例组	684	510（74.56）	174（25.44）	67.0（59.0，72.0）	678（99.12）	6（0.88）	91（13.30）	593（86.70）	448（65.50）	164（23.98）	72（10.35）
对照组	1 368	969（70.83）	399（29.17）	66.0（57.5，71.0）	1 363（99.63）	5（0.37）	172（12.57）	1 196（87.43）	744（54.39）	390（28.51）	234（17.11）
χ^2（Z）值		3.149		1.417[a]	2.239		0.218		26.396		
P 值		0.076		0.157	0.135		0.641		<0.001		

组别	饮食口味偏好［例（%）］		食物软硬偏好［例（%）］		喜食辛辣食物［例（%）］		吃饭时长［例（%）］	
	偏咸	清淡 / 适中	硬	软 / 适中	是	否	<15 min	≥ 15 min
例数	1 276		282	1 770	896	1 156	1 068	984
病例组	337（49.27）	347（50.73）	136（19.88）	548（80.12）	357（52.19）	327（47.81）	394（57.60）	290（42.40）
对照组	439（32.09）	929（67.91）	146（10.67）	1 222（89.33）	539（39.40）	829（60.60）	674（49.27）	694（50.73）
χ^2（Z）值	57.223		32.634		30.336		12.688	
P 值	<0.001		<0.001		<0.001		<0.001	

组别	规律饮食［例（%）］		食用畜肉［例（%）］		食豆制品［例（%）］		食奶制品［例（%）］		食生蒜或蒜苗［例（%）］	
	是	否	<500 g/ 周	>500 g/ 周	是	否	是	否	是	否
例数	1 694	358	1 595	457	1 737	315	1 054	998	1 235	817
病例组	515（75.29）	169（24.71）	511（74.71）	173（25.29）	563（82.31）	121（17.69）	332（48.54）	352（51.46）	358（52.34）	326（47.66）
对照组	1 179（86.18）	189（13.82）	1 084（79.24）	284（20.76）	1 174（85.82）	194（14.18）	722（52.78）	646（47.22）	877（64.11）	491（35.89）
χ^2（Z）值	37.560		5.411		4.320		3.281		26.358	
P 值	<0.001		0.020		0.038		0.070		<0.001	

组别	零食［例（%）］		饮茶［例（%）］		二手烟暴露［例（%）］		锻炼［例（%）］		发脾气［例（%）］	
	是	否	是	否	有	无	经常	无 / 偶尔	经常	无 / 偶尔
例数	662	1 390	841	1211	1 221	831	497	1 555	191	1 855
病例组	249（36.40）	435（63.60）	259（37.87）	425（62.13）	462（67.54）	222（32.46）	130（19.01）	554（80.99）	126（18.42）	558（81.58）
对照组	413（30.19）	955（69.81）	582（42.54）	786（57.46）	759（55.48）	609（44.52）	367（26.83）	1 001（73.13）	71（5.19）	1 297（94.81）
χ^2（Z）值	8.056		4.126		27.530		15.200		91.980	
P 值	0.005		0.042		<0.001		<0.001		<0.001	

组别	生闷气［例（%）］		胃部疾病史［例（%）］		一级亲属肿瘤史［例（%）］		*Hp* 感染［例（%）］		高盐饮食［例（%）］	
	经常	无 / 偶尔	有	无	有	无	阳性	阴性	是	否
例数	154	1 898	409	1 643	413	1 639	880	1 172	1 457	595
病例组	91（13.30）	593（86.70）	246（35.96）	438（64.04）	169（24.71）	515（75.29）	324（47.37）	360（52.63）	520（76.02）	164（23.98）
对照组	63（4.61）	1 305（95.39）	163（11.92）	1 205（88.08）	244（17.84）	1 124（82.16）	556（40.64）	812（59.36）	937（68.49）	431（31.51）
χ^2（Z）值	49.708		165.264		13.393		8.420		12.556	
P 值	<0.001		<0.001		<0.001		0.004		<0.001	

注：[a] 表示 Z 值，余检验统计量值为 χ^2 值；*Hp*= 幽门螺杆菌。

表 2 多因素 Logistic 回归分析变量赋值

Table 2 Variable assignment for multivariate Logistic regression analysi

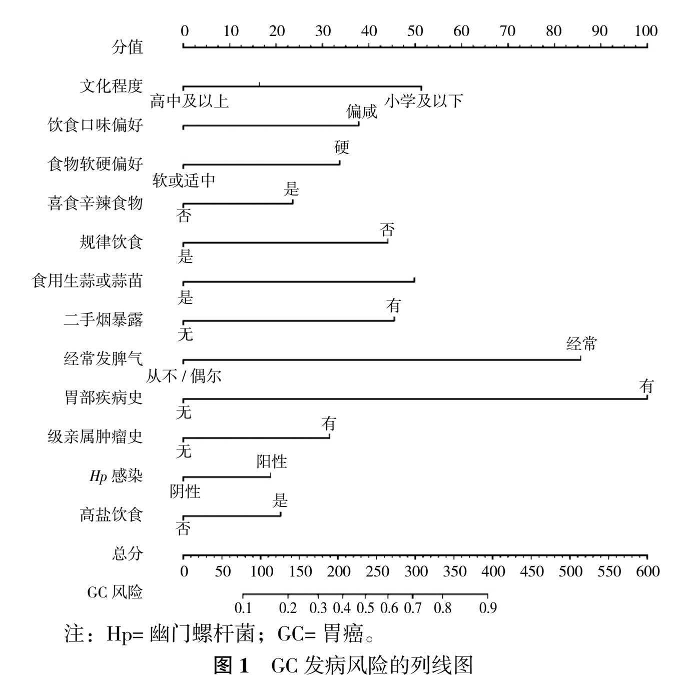
自变量	赋值
文化程度	小学及以下 =0，初中毕业 =1，高中及以上 =2
饮食口味偏好	清淡或适中 =0，偏咸 =1
食物软硬偏好	软或适中 =0，硬 =1
喜食辛辣食物	否 =0，是 =1
吃饭时长	≥ 15 min=0，<15 min=1
规律饮食	否 =0，是 =1
食畜肉	<500 g/ 周，≥ 500 g/ 周
食豆制品	否 =0，是 =1
食用生蒜或蒜苗	否 =0，是 =1
喜食零食	否 =0，是 =1
饮茶	否 =0，是 =1（每周至少一杯连续半年以上）
二手烟暴露	无 =0，有 =1
锻炼	无 / 偶尔 =0，经常 =1（>3 d ／周）
发脾气	无 / 偶尔 =0，经常 =1（>3 d ／周）
生闷气	无 / 偶尔 =0，经常 =1（>3 d ／周）
胃部疾病史	无 =0，有 =1
一级亲属肿瘤史	无 =0，有 =1
Hp 感染	阴性 =0，阳性 =1
高盐饮食	否 =0，是 =1（盐阈值 >2）

表 3 GC 发生风险的多因素 Logistic 回归分析

Table 3 Multivariate Logistic regression analysis of GC occurrence risk

变量	β	SE	Wald χ^2 值	*P* 值	OR 值（95%*CI*）
文化程度（以小学及以下为参照）					
初中毕业	-0.485	0.140	11.917	0.001	0.616（0.468~0.811）
高中及以上	-0.712	0.184	15.034	<0.001	0.491（0.342~0.703）
饮食口味偏好（以清淡 / 适中为参照）					
偏咸	0.525	0.121	18.818	<0.001	1.690（1.333~2.142）
食物软硬偏好（以软 / 适中为参照）					
硬	0.468	0.169	7.629	0.006	1.596（1.145~2.225）
喜食辛辣食物（以否为参照）					
是	0.327	0.122	7.264	0.007	1.387（1.093~1.760）
规律饮食（以否为参照）					
是	-0.612	0.149	16.938	<0.001	0.542（0.405~0.726）
食用生蒜或蒜苗（以否为参照）					
是	-0.691	0.122	32.026	<0.001	0.501（0.394~0.636）
二手烟暴露（以无为参照）					
有	0.631	0.125	25.705	<0.001	1.880（1.473~2.399）
发脾气（以无 / 偶尔为参照）					
经常	1.189	0.196	36.829	<0.001	3.283（2.236~4.819）
胃部疾病史（以无为参照）					
有	1.388	0.140	98.397	<0.001	4.008（3.046~5.273）
一级亲属肿瘤史（以无为参照）					
有	0.437	0.143	9.320	0.002	1.549（1.170~2.051）
Hp 感染（以阴性为参照）					
阳性	0.261	0.119	4.813	0.028	1.298（1.028~1.639）
高盐饮食（以否为参照）					
是	0.291	0.132	4.849	0.028	1.338（1.033~1.734）

无胃部疾病史、一级亲属无肿瘤史、无 Hp 感染、非高盐饮食者为例，总分为 357 分，GC 发病风险为 9.02%，见图 1。



注：Hp= 幽门螺杆菌；GC= 胃癌。

图 1 GC 发病风险的列线图

Figure 1 Nomogram for the risk of GC

2.4 模型验证

按照 8：2 的比例将研究对象随机分为训练集（1 641 例）和验证集（411 例）进行模型构建和验证。绘制列线图模型预测 GC 发生风险的 ROC 曲线，结果显示，训练集的 AUC 为 0.768（95%*CI*=0.744~0.792），特异度为 0.693，灵敏度为 0.702；验证集的 AUC 为 0.776（95%*CI*=0.728~0.823），特异度为 0.693，灵敏度为 0.752，见图 2。

Bootstrap 法验证结果显示，校正曲线与实际曲线一致性良好（训练集：Brier 评分 =0.177；验证集：Brier 评分 =0.176）；Hosmer-Lemeshow 检验结果显示，列线图模型的校准度良好（训练集：χ^2=4.408，*P*=0.819；验证集：χ^2=4.650，*P*=0.794），见图 3。进一步 DCA 检测结果显示，当阈值为 0.05~0.79 时，列线图模型具有较好的临床实用性，见图 4。

3 讨论

近年来，在全部恶性肿瘤的发病和死亡中，GC 居重要地位[10]。成年人口规模的增加和老龄化加速是 GC 发病率和死亡率上升的重要因素[11]，诊断不及时导致病情延误是另一关键因素。因此，亟需通过 GC 风险预测模型等进行评估和筛查，尽早识别 GC 患者并积极治疗，从而有效避免或降低 GC 发病和死亡。由于早期推行的筛查计划，近年来我国 GC 死亡率已有所下降，

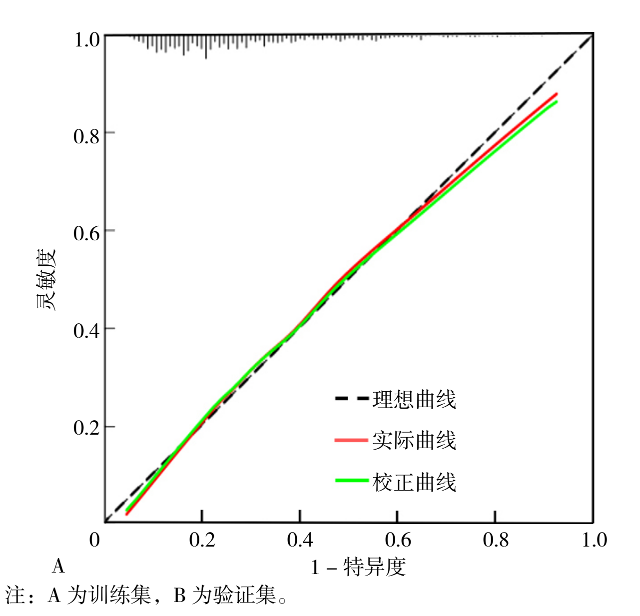
图 2 列线图模型预测 GC 发生的 ROC 曲线

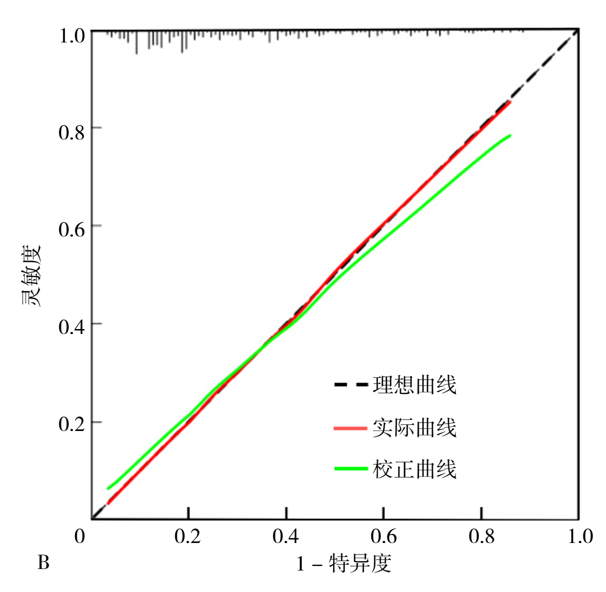
Figure 2 ROC curve of the nomogram prediction model to predict the risk of GC

但总体情况仍较差。研究指出，2013—2017 年我国 GC 筛查率为 19.62%，居常见癌种类筛查率的第 5 位，表明筛查措施尚未广泛覆盖[12]。我国 GC 防治相关指南推荐从非侵入初筛（GC 风险自评等）到侵入性精查（内镜）的 GC 筛查策略[5]。从最初的“ABC 分级法”[13]、“新 ABC 分级”[14]、CAI 等[15]开发的中国高危人群胃癌风险预测规则到胃癌前病变列线图构建等，均基于不同的预测变量构建 GC 预测模型[16]，上述研究结果促进了癌症防治领域的预测模型研究的深入发展。但当前与 GC 有关的预测模型研究多集中于 GC 治疗、治疗后感染或并发症、药物敏感性等[17-19]，而与 GC 前病变及 GC 患者早期识别相关的预测模型研究较少。因其不良预后而构成重大的全球健康挑战[19]，使 GC 防治关口前移，早期诊断和治疗是防治 GC 的关键策略，因此需不断加强和完善 GC 发生风险预测模型相关研究。

中华医学会胃癌临床诊疗指南（2021 版）[5]建议对 40 岁以上人群进行 GC 筛查。本研究针对一般人群的可改变危险因素预防，通过问卷调查分析 40 岁及以上 GC 患者与对照人群的饮食习惯、生活方式、健康状况等资料，发现文化程度、饮食口味偏好、食物软硬偏好、辛辣食物偏好、规律饮食、食用生蒜或蒜苗、二手烟暴露、经常发脾气、胃部疾病史、一级亲属肿瘤史、Hp 感染及高盐饮食是 GC 发病的独立影响因素。较高学历（OR=0.491）是 GC 的保护因素，这与 HUANG 等[20]的研究结果一致。因为受教育程度高者普遍具有较高的健康素养水平，其对健康信息的理解、甄别和应用能力更高[21]，更易采取健康的生活方式，如：均衡饮食、适量运动、定期体检等，从而降低患 GC 的风险。

饮食习惯分析结果显示口味偏咸（OR=1.690）、高盐饮食（OR=1.338）、食物偏硬（OR=1.596）、喜食辣食（OR=1.387）与 GC 发生风险呈正相关，而规律饮食（OR=0.542）、食用生蒜或蒜苗（OR=0.501）与胃癌发生风险呈负相关，该结果与其他研究一致[22-24]。高盐饮食是 GC 最重要的危险因素之一，WHO 建议每日摄入量低于 5 g，过量摄入会刺激胃粘膜，可能导致黏膜萎缩、DNA 合成增加和细胞增殖等，从而增加 GC 患病风险[25]。干硬或辛辣食物在消化过程中也会对胃粘膜产生物理或化学刺激，增加胃炎等胃部疾病风险[10]。此外，饮食不规律可导致营养不良、肥胖、消化道疾病等[5]，还会诱发胃粘膜反复损伤，损害其保护机制，长此以往可导致 GC 的发展[26]。本研究结果显示食用

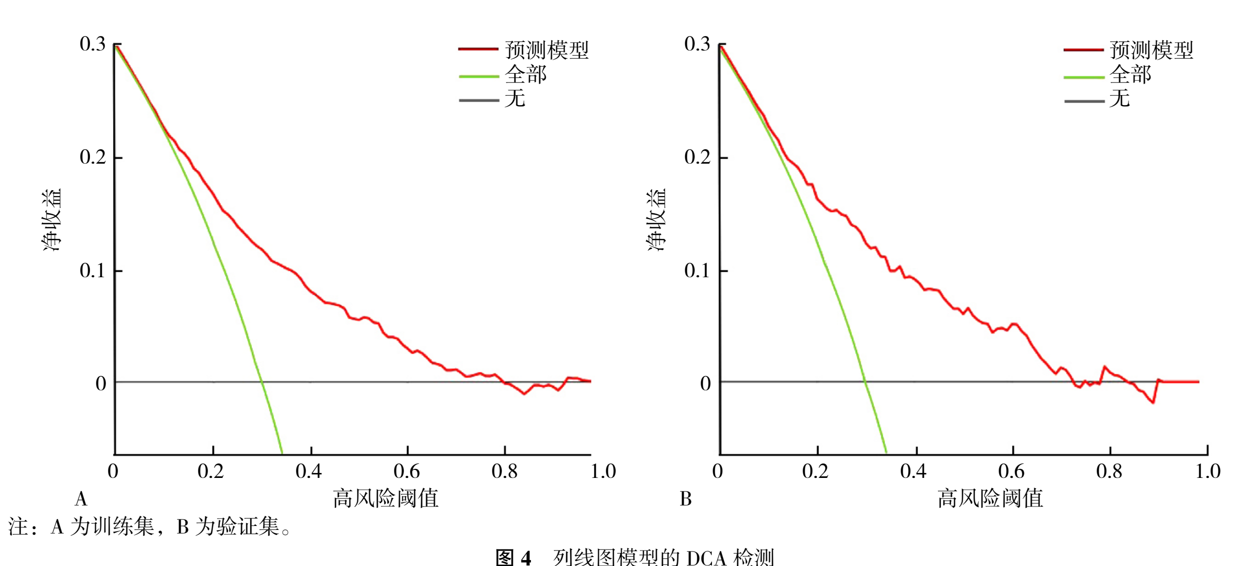




注：A 为训练集，B 为验证集。

图 3 列线图模型的校准曲线

Figure 3 Calibration curve of the Nomogram model



注：A 为训练集，B 为验证集。

图 4 列线图模型的 DCA 检测

Figure 4 DCA analysis of the Nomogram model

生蒜或蒜苗可降低 GC 发生风险，与 Li 等[27]的研究结果一致。补充大蒜（提取物和油）可降低GC发病和死亡，因为蒜及其衍生物具有抗氧化、抗菌和免疫调节的特性，同时大蒜富含烯丙基衍生物，可通过调节致癌物的代谢酶活性、防止 DNA 加合物形成、改变细胞周期等抑制细胞增殖和肿瘤生长[28]。

生活方式分析结果显示二手烟暴露（*OR*=1.880）和经常发脾气（*OR*=3.283）是 GC 发生的危险因素，这与美国一项纵向研究结果一致[29]。研究表明，烟草产品的烟雾中含有大量的有害物质和致癌物，二手烟暴露者吸入烟雾中的尼古丁、焦油等有害物质和致癌物是吸烟者的 5.1 倍，长期高水平的二手烟暴露会显著增加肺癌、GC、结直肠癌等多种恶性肿瘤患病风险[30]。此外，本研究结果表明经常发脾气是 GC 的独立危险因素，这与熊文燕等[31]、沙宇婷等[32]的研究结论一致。这可能是因为发脾气会影响胃肠功能，导致消化系统紊乱，因而加重胃肠道负担，造成消化不良、胃黏膜损伤等问题，经常发脾气而长期处于消化系统紊乱和胃黏膜损伤状态，从而增加 GC 的发病风险[33]。

本研究显示胃部疾病史（*OR*=4.008）、*Hp* 感染（*OR*=1.298）是 GC 发生的独立危险因素，这与孙凯旋等[34]的研究结果一致。*Hp* 感染是胃癌最重要的致病因素，其通过分泌特异的毒力因子与生活环境、宿主遗传及胃内微生态共同作用致癌[6]。*Hp* 还与慢性胃炎、消化性溃疡等一些严重的胃肠道疾病的发生密切相关[35]。根据 CORREA 等[36]提出的 GC 进展级联模型，从胃黏膜炎症反应开始，急性胃炎逐渐转化为慢性胃炎，随后在长期的炎症过程中，胃黏膜反复损伤和修复，进而发展为萎缩性胃炎、肠上皮化生或异型增生等胃部疾病，最终导致 GC 的发生。此外，研究发现一级亲属肿瘤史（*OR*=1.549）是 GC 的危险因素，与既往研究结果一致[22, 32]。其原因可能在于癌症家族史不仅会增加个体患 GC 的风险，还会增加胃粘膜萎缩的风险[37]，从而使患 GC 的风险增加。

本研究基于筛选出的 GC 危险因素构建列线图模型，结果显示模型具有良好的判别能力，训练集 AUC 为 0.768，验证集 AUC 为 0.776。校准曲线和 Hosmer-Lemeshow 拟合优度检验显示训练集和验证集的预测概率与实际概率一致性较好。DCA 曲线显示模型具有良好的临床效用，表明使用该列线图模型预测 GC 发生风险并进行相应的干预可以获得正向净收益。本研究构建的列线图模型具有较好的预测能力，可准确识别个体的 GC 发病风险，为 GC 早期防治提供定量、直观的参考工具。

本研究还存在一些不足。（1）为控制选择偏倚，本研究按年龄和性别匹配对照人群，导致无法评估年龄和性别在GC发生风险预测中的作用；（2）病例数量较少，统计效力较弱，导致研究结果外推受限；（3）GC 是一个病程较长的复杂疾病，期间暴露因素可能随时间发生变化，可能影响本研究中暴露因素与疾病之间的关联强度，导致所得效应值与其他研究结论稍有出入；（4）本研究数据来源较为单一，未利用外部数据进一步验证模型效能，在后续研究中将借助其他地区数据对该模型进行外部验证，以提高模型的预测精度和通用性。

综上所述，本研究发现文化程度较低、饮食口味偏咸、食物偏硬、喜食辣食、不规律饮食、二手烟暴露、经常发脾气、胃部疾病史、一级亲属肿瘤史、Hp 感染及高盐饮食是 GC 发生的危险因素；基于此结果构建的列线图模型能有效预测 GC 发生风险。该模型纳入 GC 可改变危险因素进行预测，属非侵入性自我诊断方法，

经济成本低易被筛查人群所接受，有助于早期识别GC高危人群，并通过进一步胃镜检查等手段早诊早治，提高治疗效果和生存率；也可通过危险因素评估可能发生GC的风险，从而增强健康意识、提高生活质量，进而改变不良生活习惯，使GC防治关口前移。但该模型是基于单中心研究获得，尚需多中心、前瞻性、大样本研究进一步验证。

作者贡献：周倩提出主要研究目标，负责研究的构思与设计，撰写论文；周倩、吴晓敏、王宝华进行数据的收集与整理，统计学处理，图、表的绘制与展示；周倩、严若菡、蔚苗进行论文的修订；王宝华、吴静负责文章的质量控制与审查，对文章整体负责，监督管理。

本文无利益冲突。

参考文献

[1] SUNG H，FERLAY J，SIEGEL R L，et al. Global cancer statistics 2020：GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin，2021，71(3)：209-249. DOI：10.3322/caac.21660.

[2] 闫超，陕飞，李子禹. 2020年全球胃癌负担分析：聚焦中国流行现状[J]. 中国肿瘤，2023，32（3）：161-170.

[3] CHOI S，SONG J H，LEE S J，et al. Trends in clinical outcomes and long-term survival after robotic gastrectomy for gastric cancer：a single high-volume center experience of consecutive 2000 patients[J]. Gastric Cancer，2022，25（1）：275-286. DOI：10.1007/s10120-021-01231-3.

[4] 中国抗癌协会胃癌专业委员会，中国医师协会外科医师分会上消化道外科医师委员会，中国人群健康风险管理协作组-胃癌专业组. 中国人群胃癌风险管理公众指南（2023版）[J]. 中华医学杂志，2023，103（36）：2837-2849. DOI：10.3760/cma.j.cn112137-20230608-00968.

[5] 中华医学会肿瘤学分会，中华医学会杂志社. 中华医学会胃癌临床诊疗指南（2021版）[J]. 中华医学杂志，2022，102（16）：1169-1189. DOI：10.3760/cma.j.cn112137-20220127-00197.

[6] KARPINSKI M J，HÜSING J，CLAASSEN K，et al. Combining PSMA-PET and PROMISE to re-define disease stage and risk in patients with prostate cancer：a multicentre retrospective study[J]. Lancet Oncol，2024，25（9）：1188-1201. DOI：10.1016/S1470-2045（24）00326-7.

[7] HARRISON H，PENNELLS L，WOOD A，et al. Validation and public health modelling of risk prediction models for kidney cancer using the UK Biobank[J]. BJU Int，2022，129（4）：498-511. DOI：10.1111/bju.15598.

[8] POOROLAJAL J，MORADI L，MOHAMMADI Y，et al. Risk factors for stomach cancer：a systematic review and meta-analysis[J]. Epidemiol Health，2020，42：e2020004. DOI：10.4178/epih.e2020004.

[9] ZHOU X Z，LYU N H，ZHU H Y，et al. Large-scale，national，family-based epidemiological study on Helicobacter pylori infection in China：the time to change practice for related disease prevention[J]. Gut，2023，72（5）：855-869. DOI：10.1136/gutjnl-2022-328965.

[10] HAN B F，ZHENG R S，ZENG H M，et al. Cancer incidence and mortality in China，2022[J]. J Natl Cancer Cent，2024，4（1）：47-53. DOI：10.1016/j.jncc.2024.01.006.

[11] XIA C F，DONG X S，LI H，et al. Cancer statistics in China and United States，2022：profiles，trends，and determinants[J]. Chin Med J，2022，135（5）：584-590. DOI：10.1097/CM9.0000000000002108.

[12] 陈万青，李霓，曹毛毛，等. 2013—2017年中国城市癌症早诊早治项目基线结果分析[J]. 中国肿瘤，2020，29（1）：1-6.

[13] FURIHATA C. Human gastric cancer risk screening：from rat pepsinogen studies to the ABC method[J]. Proc Jpn Acad Ser B Phys Biol Sci，2021，97（8）：462-478. DOI：10.2183/pjab.97.023.

[14] WANG X T，JI Z Z，HAN F，et al. A comparative study of new gastric cancer screening scoring system and new ABC method for screening gastric cancer and precancerous lesions[J]. Zhonghua Nei Ke Za Zhi，2021，60（3）：227-232. DOI：10.3760/cma.j.cn112138-20200512-00476.

[15] CAI Q C，ZHU C P，YUAN Y，et al. Development and validation of a prediction rule for estimating gastric cancer risk in the Chinese high-risk population：a nationwide multicentre study[J]. Gut，2019，68（9）：1576-1587. DOI：10.1136/gutjnl-2018-317556.

[16] SHI C E，TAO R，WANG W S，et al. Development and validation of a nomogram for obesity and related factors to detect gastric precancerous lesions in the Chinese population：a retrospective cohort study[J]. Front Oncol，2024，14：1419845. DOI：10.3389/fonc.2024.1419845.

[17] GAO P，XIAO Q，TAN H，et al. Interpretable multi-modal artificial intelligence model for predicting gastric cancer response to neoadjuvant chemotherapy[J]. Cell Rep Med，2024：101848. DOI：10.1016/j.xcrm.2024.101848.

[18] 马兴好，江晓阳，王传霞，等. 预测胃癌D2根治术后感染风险列线图模型的建立与验证[J]. 中国感染控制杂志，2023，22（9）：1013-1020. DOI：10.12138/j.issn.1671-9638.20234083.

[19] LI X X，WANG X X，YU F X，et al. Development and validation of a prognostic and drug sensitivity model for gastric cancer utilizing telomere-related genes[J]. Transl Oncol，2024，52：102232. DOI：10.1016/j.tranon.2024.102232.

[20] HUANG L，CHEN L，GUI Z X，et al. Preventable lifestyle and eating habits associated with gastric adenocarcinoma：a case-control study[J]. J Cancer，2020，11（5）：1231-1239. DOI：10.7150/jca.39023.

[21] 张刚，李英华，李莉，等. 2021年我国城乡居民健康素养水平及其影响因素研究[J]. 中国健康教育，2024，40（5）：387-391，400. DOI：10.16168/j.cnki.issn.1002-9982.2024.05.001.

[22] 朱建梅，何科茗. 胃癌风险人群筛查及危险因素分析[J]. 浙江创伤外科，2023，28（4）：623-626，630. DOI：10.3969/j.issn.1009-7147.2023.04.004.

[23] KANG K，BAGAOISAN M A，ZHANG Y X. Unveiling the younger face of gastric cancer：a comprehensive review of epidemiology，

risk factors, and prevention strategies [J]. Cureus, 2024, 16(6): e62826. DOI: 10.7759/cureus.62826.

[24] WU X M, CHEN L L, CHENG J X, et al. Effect of dietary salt intake on risk of gastric cancer: a systematic review and meta-analysis of case-control studies [J]. Nutrients, 2022, 14(20): 4260. DOI: 10.3390/nu14204260.

[25] FANG X X, WEI J Y, HE X Y, et al. Landscape of dietary factors associated with risk of gastric cancer: a systematic review and dose-response meta-analysis of prospective cohort studies [J]. Eur J Cancer, 2015, 51(18): 2820-2832. DOI: 10.1016/j.ejca.2015.09.010.

[26]《中国吸烟危害健康报告》编写组.《中国吸烟危害健康报告2020》概要 [J]. 中国循环杂志, 2021, 36(10): 937-952. DOI: 10.3969/j.issn.1000-3614.2021.10.001.

[27] LI W Q, ZHANG J Y, MA J L, et al. Effects of Helicobacter pylori treatment and vitamin and garlic supplementation on gastric cancer incidence and mortality: follow-up of a randomized intervention trial [J]. BMJ, 2019, 366: l5016. DOI: 10.1136/bmj.l5016.

[28] WANG Y Y, HUANG P, WU Y F, et al. Association and mechanism of garlic consumption with gastrointestinal cancer risk: a systematic review and meta-analysis [J]. Oncol Lett, 2022, 23(4): 125. DOI: 10.3892/ol.2022.13245.

[29] ZHANG Z M, LI Z P, ZHANG X Y, et al. Association between secondhand smoke and cancers in adults in the US population [J]. J Cancer Res Clin Oncol, 2023, 149(7): 3447-3455. DOI: 10.1007/s00432-022-04266-w.

[30] 国家卫生健康委员会. 中国吸烟危害健康报告. 2020[M]. 北京: 人民卫生出版社, 2021: 109-114.

[31] 熊文艳, 范义兵, 冯小武, 等. 2018—2019年南昌市城市居民胃癌前疾病危险因素病例对照研究 [J]. 中华肿瘤防治杂志, 2021, 28(24): 1856-1861. DOI: 10.16073/j.cnki.cjcpt.2021.24.03.

[32] 沙宇婷, 王雪威, 王宝华. 中国南方胃癌危险因素病例-对照研究Meta分析 [J]. 中华肿瘤防治杂志, 2020, 27(15): 1269-1276. DOI: 10.16073/j.cnki.cjcpt.2020.15.15.

[33] VAHID F, RAHMANI W, DAVOODI S H, et al. Mental health conditions, including depression and stress, are associated with increased odds of gastric cancer-insights into the role of diet: a case-control study [J]. Nutrients, 2023, 15(23): 4981. DOI: 10.3390/nu15234981.

[34] 孙凯旋, 张眉佳, 廖成功, 等. 中国人群胃癌发病影响因素的Meta分析 [J]. 现代肿瘤医学, 2022, 30(8): 1431-1438.

[35] CHEOK Y Y, LEE C Y Q, CHEONG H C, et al. An overview of Helicobacter pylori survival tactics in the hostile human stomach environment [J]. Microorganisms, 2021, 9(12): 2502. DOI: 10.3390/microorganisms9122502.

[36] CORREA P, HAENSZEL W, CUELLO C, et al. A model for gastric cancer epidemiology [J]. Lancet, 1975, 2(7924): 58-60. DOI: 10.1016/s0140-6736(75)90498-5.

[37] 梁国英, 曲智慧, 李庆伟. 慢性萎缩性胃炎致病因素的中西医研究进展 [J]. 中国中西医结合消化杂志, 2022, 30(5): 378-382.

(收稿日期：2024-08-29；修回日期：2024-11-11)

(本文编辑：李卫霞)



附表

附表 1 4 省 14 个县（区）研究对象纳入数量（例）

Table S1 Number of subjects included from 14 counties （districts） in 14 provinces.

组别	江苏				安徽			
	盱眙县	金湖县	射阳县	颍东区	埇桥区	灵璧县	蒙城县	寿县
病例组	46	50	49	54	46	42	48	49
对照组	96	96	107	96	89	81	101	103
合计	142	146	156	150	135	123	149	152

组别	山东		河南			
	汶上县	巨野县	扶沟县	西平县	罗山县	沈丘县
病例组	49	47	63	51	53	50
对照组	105	92	88	98	106	97
合计	154	139	151	149	159	147